

A Cultural Heritage Application Schema: Achieving Interoperability of Cultural Heritage Data in INSPIRE*

Carlos Fernández-Freire¹, Isabel del-Bosque-González¹, Juan Manuel Vicent-García², Esther Pérez-Asensio¹, Alfonso Fraguas-Bravo², Antonio Uriarte-González², Pastor Fábrega-Álvarez³ and César Parcero-Oubiña³

¹GIS Laboratory (CCHS) – Spanish National Research Council (CSIC), Madrid, Spain; carlos.ffreire@cchs.csic.es; isabel.delbosque@cchs.csic.es; esther.perez@cchs.csic.es;

²Institute of History (IH-CCHS) – Spanish National Research Council (CSIC), Madrid, Spain; juan.vicent@cchs.csic.es; alfonso.fraguas@cchs.csic.es; antonio.uriarte@cchs.csic.es

³Institute of Heritage Sciences (Incipit), Spanish National Research Council (CSIC), Santiago de Compostela, Spain; pastor.fabrega-alvarez@incipit.csic.es; cesar.parcero-oubiña@incipit.csic.es

Abstract

Cultural Heritage is a group of resources identified as a reflection of certain values, beliefs, knowledge and tradition by a certain community. These resources are subject to management, preservation and diffusion through legislative and administrative means, which makes cultural heritage fall within the scope of *Protected sites*, one of the spatial data themes established in Annex I of the INSPIRE Directive. The INSPIRE Data Specification on Protected Sites thus serves as the starting point for modelling cultural heritage information in order to implement, distribute and share it in an interoperable framework based on Spatial Data Infrastructures. Unfortunately, this data specification was primarily conceived for

*This work is licensed under the Creative Commons Attribution-Non commercial Works 3.0 License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/> or send a letter to Creative Commons, 543 Howard Street, 5th Floor, San Francisco, California, 94105, USA.

natural protected sites, which makes an extension necessary if it is to be applied to cultural features. This paper proposes an extension composed of three parts: one devoted to administrative information—including legal protection—, another describes the feature itself, and, a third part is dedicated to the inclusion of additional documentation (texts, images, etc.).

Keywords: Cultural Heritage, INSPIRE, Protected Sites, Conceptual Data Model, Application Schema, Interoperability.

1. INTRODUCTION

This paper presents an application schema for cultural heritage information within INSPIRE. The proposed schema addresses the need to enhance interoperability in relation to the field of Cultural Heritage and in relation to INSPIRE integration within the Data Specification on Protected Sites, one of the reference themes included in the Annex I of the Directive. Existing conceptual formalizations of Cultural Heritage do not suit the INSPIRE requirements, so a specific schema is proposed in this paper, as an extension of the Protected Sites Full schema.

The Cultural Heritage Application Schema is constituted by three parts, a legal-administrative part, rooted on the concept of protected sites, as defined for managing and preserving cultural features; a strictly cultural part, centred on the real-world entities that are the object of protection; and a documentary part regarding information resources related to cultural entities. This approach will surely improve the dissemination and exploitation of cultural heritage information.

1.1. Cultural Heritage as Protected Sites

“Cultural heritage is a group of resources inherited from the past which people identify, independently of ownership, as a reflection and expression of their constantly evolving values, beliefs, knowledge and traditions” (Faro, 2005: Art. 2.a). It is a complex reality, based on the importance of perceived value (about this concept, see for example Querol, 2010). Cultural heritage is usually considered to include both *tangible/material* and *intangible/immaterial* elements. Within the first category there are two basic types of elements: *immovable* or *built*, which are intrinsically linked to their spatial location—a church, a bridge or an archaeological site—and *movable*, which can be transported without losing their main intrinsic characteristics—a painting, a piece of pottery or a sword—. On the other hand, the so-called intangible heritage comprises all those human activities whose existence is not immanent but depends on their performance, being closely linked to memories and folklore (e.g. a dance, a pilgrimage or a traditional kitchen recipe).

One of the main concerns regarding cultural heritage is its protection and preservation, which has led to the worldwide creation of legal systems devoted to this task at different administrative levels: international, national, regional and local. This process is closely related to the emergence of modern democratic states, where public administrations have undertaken the safeguard of cultural heritage, and guaranteeing the rights of citizens to access it.

Much cultural heritage consists of things that can be found in a place, or linked to a geographical location. This is especially evident in the case of immovable features (buildings, historical places, archaeological sites, cultural landscapes, etc.), whose integrity is intrinsically tied to the place they occupy. As for movable items, their heritage value may subsist regardless of its location, though it is frequently lessened when moved, since such value depends to a great extent on the successive spatial contexts where it has been through its life cycle (from the places where these movable features were created and used to the museums and collections where they are preserved). Finally, intangible heritage will always be linked to its spatial dimension, as the context for the human activities which generate and recreate it, though its heritage value remains initially detached from any specific object-based appearance. The spatial dimension of cultural heritage is central to understand its nature and to ensure the protection commitments made on behalf of public administrations.

This geographic nature and the preservation obligations make cultural heritage a reality that has to be included within territorial planning policies. This becomes especially relevant when heritage preservation conflicts with landscape change brought about by development, a challenge which has attained an unprecedented scale (Vicent García, 2007). This and the problems that arise when guaranteeing accessibility are the main reasons for requiring public intervention to ensure an adequate protection of cultural heritage.

The implementation of protective measures usually involves the administrative identification and catalogue of cultural heritage features. This has created myriad datasets (catalogues, inventories), each one managed with different criteria, reflecting the multiple normative frameworks supported by the corresponding administrations at various levels (local, regional, national, international). In order to gather and manage heritage information, administrations have applied, with different degrees of complexity and efficiency, Information and Communication Technologies (ICTs), including Geographical Information Technologies (GITs) when reflecting spatial information. These applications have usually been focused on management tasks, for their use on behalf of heritage technicians. In recent years, there has been increased awareness of the need to make this information available for society, researchers, administrators, stakeholders, etc. In some cases this has meant the distribution of georeferenced heritage data via

SDIs: in Spain there is the architectural heritage WMS of Navarra or the heritage WMS of Mallorca, among others.

These worthy efforts notwithstanding, the distribution of cultural heritage information is still heterogeneous, with multiple and varied cultural heritage datasets, which are often fragmentary, unconnected and barely accessible (Corns and Shaw, 2010: 2). An interoperability framework for sharing and integrating cultural heritage information from different datasets and also for linking them with other layers of geographical information (environmental information, infrastructures, urban areas, etc.) is therefore needed. The Protected Sites (PS) Data Specification includes protected features both cultural and natural, which are relevant to territorial planning, and hence have to be provided within the INSPIRE Directive.

1.2. Cultural Heritage Data Fit Within Inspire

The INSPIRE Directive (2007/2/EC) is intended to solve problems “regarding the availability, quality, organisation, accessibility and sharing of spatial information”, guaranteeing the interoperability of spatial datasets through network services across Europe (2007/2/EC, par.2). Cultural heritage features are included in the INSPIRE Directive, and as long as they are protected sites, they are considered relevant information, included as one of the Annex I themes.

Protected sites are defined as “an area designated or managed within a framework of international, Community and Member States’ legislation to achieve specific conservation objectives” (INSPIRE Thematic Working Group Protected sites, 2010). There are, nonetheless, two main data specifications inside INSPIRE that might concern cultural heritage data: the Data Specification on Protected Sites (PS Data Specification), that develops the 9th theme of the Annex I, and the Data Specification on Area management/restriction/regulation zones and reporting units (AM Data Specification), that develops the 11th theme of Annex III. These data specifications include the technical documentation of the application schema, and a conceptual schema —expressed in UML— that defines the content and structure of the data required by one or more applications, thus guaranteeing its correct understanding.

The PS Data Specification applies to all protected sites defined by international, EU or national legislation, even if the legislation is managed at the local, provincial or regional level. It extends the definition of *Protected site* given by the INSPIRE Directive by appealing to that of the International Union for Conservation of Nature (IUCN): “an area of land and/or sea especially dedicated to the protection and maintenance of biological diversity, and of natural and associated cultural resources, and managed through legal or other effective means” (INSPIRE Thematic Working Group Protected sites, 2010: 1). According to this data specification the objectives for protection may include, among others, the protection of

person-made objects including buildings, archaeological sites and other cultural objects.

The AM Data Specification covers a wide range designed to describe “zones established in accordance with specific legislative requirements to deliver specific environmental objectives related to any domain, such as air water, soil, biota (plants and animals), natural resources, land and land use”. It may therefore overlap with *Protected sites*, in which case the AM Data Specification is to be used. It establishes a clear distinction for these overlapping cases: Protected sites are intended to manage, regulate and restrict activities to conserve nature, biodiversity and cultural heritage *exclusively*. When they are established to deliver multiple objectives, they should be made available as Area management / restriction / regulation zones objects.

Cultural heritage also overlaps with the Data Specification on Buildings, for architectural heritage is to be modelled as constructed buildings under that data specification. Nonetheless, many heritage features do not fit within the definition of a building provided by it. The Data Specification on Buildings will collect the real geometry of the building, that constitutes relevant information, leaving the geometry of the protected area apart.

The scope of cultural heritage data falls entirely within the definition of *Protected sites*, although their preservation in the face of development would require management measures modelled under the AM Data Specification, as it may be the case of controlling urban or industrial developments. Nevertheless, both data specifications include information regarding the legal conditions of areas protecting a cultural heritage site, but the actual feature being protected is omitted. This poses two different problems: the absence of a definition which can enable the incorporation of the legal specificities of cultural protected sites, and the characterization of a common framework to convey a minimum amount of data regarding the nature of the cultural heritage protected, which has a spatial dimension as well. The latter has been widely addressed under Annex III themes like *Habitats and Biotopes*, *Species Distribution* or *Buildings*, all of them linked to protected sites, but the specificity and diversity of cultural heritage sites is not to be found within any Annex III theme.

The Cultural Heritage Application Schema proposed in this paper suggests the inclusion of cultural heritage protected areas as a special kind of protected sites and develops a minimal schema to include information about the real-world protected entity. This approach involves some difficulties because, although culturally valued places are explicitly included in those definitions of what an INSPIRE protected site is, the corresponding data specification has been modelled on the particulars of natural protected sites. Overcoming these difficulties shall lead to the integration of legal aspects concerning cultural heritage protected sites but,

as it has been pointed out, the PS Data Specification offers a narrow framework for protected sites as statutory designations (McKeague, Corns and Shaw, 2012), centred mainly on attributes regarding legal aspects. The protected entity is incorporated as an attribute of the legal entity, using the same structure that the specifications that collect information on equivalent features: *Habitats and Biotopes* and *Species Distribution*.

2. FORMALIZATIONS OF CULTURAL HERITAGE DATA

As far as we know there are no previous proposals for technically integrating cultural heritage information within the INSPIRE framework (but see McKeague et al, 2012; McKeague, 2012). Although work has been done for the creation of heritage-related SDIs (e.g. Omtzigt et al, 2008; Lage et al, 2009; McKeague, Corns, and Shaw, 2012; see also Parcero-Oubiña, 2012), the conceptual or data models behind those experiences are usually not shown in detail. On the other hand, a number of proposals exist that provide either a general conceptual modelling of cultural heritage elements, well beyond its spatial dimension, or spatially-oriented data models for specific themes. They are not designed, however, to meet the INSPIRE requirements, although they may provide essential theoretical and conceptual foundations for the development of a new proposals.

A number of specific spatial data models for concrete thematic areas of the wider field of cultural heritage have been developed so far. Perhaps archaeology is one of the heritage-related disciplines where a greater consciousness of the need for the management of spatial data exists. As a result, detailed spatial data models exist, for instance, for the documentation of information about archaeological excavations (e.g. Meyer et al, 2007; Pfoser et al, 2007; Katsianis et al, 2008). Although partial in scope and too detailed for our purpose here, they provide a sound definition of concepts and relationships which facilitate their integration as extensions of the general model provided by the Cultural Heritage Application Schema.

With regard to models that engage cultural heritage as a whole, the CIDOC Conceptual Reference Model –CIDOC-CRM– (Doerr, 2003; Crofts et al, 2010) is the best known and widespread proposal. CIDOC-CRM constitutes a major effort to build a comprehensive reasoning tool to describe many processes related to the description of heritage features, the documentation of the circumstances involved in their creation and life cycle, and the management of their present day condition. Developed in the context of the ICOM since the 1990s, it is especially aimed at “the curated knowledge of museums” and the management of “all information required for the exchange and integration of heterogeneous scientific documentation of museum collections” (Crofts et al, 2010: i–ii). CIDOC-CRM has attained the status of ISO standard (ISO 21127:2006) and has become a major tool for the integration of information in some fields within cultural heritage practice, as

diverse as data from archaeological repositories (Tudhope et al, 2011) or ethnomusicology (Strle and Marolt, 2012), among many others¹.

However, when considered from the perspective of geospatial information, some drawbacks exist to consider it as a direct source to build an INSPIRE-compliant data model. First, and foremost, CIDOC-CRM is not a data model and it hardly regards the spatial dimension, but rather it is a conceptual reference model, an ontology. It provides a framework of reference to allow the integration and interoperability of different datasets, but it does not define a data model itself (for instance, classes do not have attributes). As we will see later on, some of the capabilities it provides are useful, and some CIDOC-CRM concepts allow data in the Cultural Heritage Application Schema to be easily mapped in its terms, as has been done in other projects where the management of heritage-related spatial information was needed (e.g. Willmes et al, 2012).

However, besides that, two additional questions should be remarked. Firstly, CIDOC-CRM is primarily aimed at objects, rather than at geographical elements. This is not to say that CIDOC-CRM assumes that objects do not have a spatial dimension, or that their spatial location is not considered. Actually, it has developed a thorough and robust model to describe the spatial location of things (Dorr, 2003, 87–88; Crofts et al, 2010: xix). However, this model is not primarily aimed at geographical representation or at the management of spatial objects, but at spatial reasoning: things and their spatial location are modelled as two separate classes. The key class there, *Place*, is defined as an abstract location, “in the pure sense of physics: independent from temporal phenomena and matter” (Crofts et al, 2010: 22). Geometry is not conceived as a property of things, but as one of the possible ways to describe a *Place*, which in its turn is a spatial location where things or objects can be found. This is an extremely rich and powerful model that allows, for instance, to easily describe changes in the location of things through time, which is very coherent with the primary concern that CIDOC-CRM has with objects and with its event-based character (Janowicz, 2009). Besides that, it allows simple management of relative locations, which are considered to be “often more relevant in the context of cultural documentation and tend to be more precise” (Crofts et al, 2010: 22) than geospatial locations. Albeit as rich, coherent and powerful as this is, it is a totally different approach with regards to INSPIRE, where location and geometry is an inherent property of things. Examples of the integration of GIS spatial data with the CIDOC-CRM model result, as would be expected, in a data structure where things and spatial entities are two different classes of objects (e.g. Cripps et al, 2004; Hiebel, Hanke, and Hayek, 2013).

¹ A list of cases in <http://www.cidoc-crm.org/references.html>

MIDAS is another proposal worth mentioning here. It is a data standard developed under the coordination of English Heritage, “for information about the historic environment” in England (Heritage, 2012: 8). The MIDAS approach has been adopted and extended beyond England, as in the European project CARARE (Papatheodorou et al, 2011). MIDAS is not an abstract ontology, but a specific data model to document and describe different types of heritage elements. Any of the *Heritage Assets* it defines can be represented as spatial objects, by using what is called a *Map Depiction*. However, since this is not a spatially-oriented system, (1) location is not mandatory and (2) the digital nature of map representations is also optional: “The use of GIS is highly recommended. It provides additional options for recording spatial extents as well as extended querying functionality” (Heritage, 2012: 72). *Map Depiction* is, therefore, a separate class, instead of an embedded component of objects.

The interesting point here, however, is the conceptual difference it establishes between the three types of possible Heritage Assets used to describe spatial entities: *Area*, *Monument* and *Artefact and ecofact*. These concepts set apart what can be described as real-world things (monuments and objects) and other heritage-relevant spatial locations, resulting from research or management processes and decisions. *Area* is defined as “a defined area of land, urban or seascape, of significance for an understanding of the historic environment and its management” (p. 25), and examples include areas covered by a research project, designated areas of protection, characterization areas, etc. This is very much in coincidence with the approach of the Cultural Heritage Application Schema, since it avoids the usual confusion when managing heritage elements between the description (and geometry) of designated areas and of the elements (features, objects) that are protected by those designations.

Such a common entanglement is found, for instance, in the CIDOC International Core Data Standard for Archaeological Sites and Monuments (Thornes and Bold, 1998). Although not aimed at the use of spatial objects, location is here mandatory information that may be expressed in different ways, being geospatial location one of them. However, protection is conceived as one of the possible properties of monuments and sites, and so location is unique for both (for instance, a monument and its corresponding protected place will share the same location), which could easily cause redundancies or topological problems.

It could be argued that, not being focused on the management of spatial objects, topological questions are not a primary issue for the Core Data Standard. Nonetheless, it is only expected that databases developed in the realm of cultural resources management (CRM) reproduce the confusion between protected places and things protected. A typical CRM approach is that geometry corresponds to a protected demarcation, while descriptive attributes describe the real-world features within. A well-known case for us is that of the otherwise excellent infor-

mation system of the Instituto Andaluz de Patrimonio Histórico (IAPH, 2011). This, as we shall see later on, causes the confusing effect that changes in attributes of the cultural entity (including geometry) mean changes in attributes of the legal entity of protection.

The evaluation of the experiences mentioned as well as others alike led to the definition of the data model that is described in the following sections. Well established concepts and widespread reference models, such as CIDOC-CRM, have been especially considered, so that classes in the Cultural Heritage Application Schema were designed to be easily mapped against them if needed. A key difference, however, exists between the INSPIRE approach and most of the others, since it conceives the spatial dimension of things (location and geometry) as something inherent to them and embedded as one of their properties, rather than as a separate concept such as *Place* in CIDOC-CRM and other models.

3. INTEGRATING CULTURAL HERITAGE DATA IN *PROTECTED SITES*

Our proposal intends to provide an application schema that will offer a generic, extendable and interoperable framework for the development of cultural heritage SDIs. It is designed to be an extension of the PS Data Specification, in order to enable a full integration of cultural heritage data.

3.1. The Data Specification on Protected Sites

The INSPIRE PS Data Specification offers two application schemas, Simple and Full. In both schemas the *ProtectedSite* class contains attributes regarding the legal figure of protection (Table 1). The simple schema holds a very limited set of fundamental attributes that establishes the core specification (geometry and identifier, geographical name, etc.), while the full one includes additional *voidable* attributes (site description, spatial resolution, ownership, legal expiry date, etc.), that must be provided if included in the dataset, unless they do not exist in the real world (*unpopulated*) or the correct value is not known by the data provider (*unknown*).

Table 1: Attributes of Protected Sites

Legal aspects referring to the document that enable its protection , such as the date, a document citation, an expiry date, the institution responsible for its protection, the type of protected site according to a predefined schema, the type of protected site according to the nature of the entity, an ownership indication and information on the official size of the protected site.
Aspects related to their geographical nature : the site name, that refers to the geographical name and has to be filled according to the INSPIRE Data Specification on Geographical Names (2009); the spatial resolution; and its geometry.

<p>The real-world entities protected by a designated protected site, more fully developed under other INSPIRE Annex III themes. As it has been already pointed out, Annex III does not include any theme suitable for the incorporation of cultural heritage data.</p>
<p>Attributes relevant only to natural protected sites that may be filled in for some heritage places. Nonetheless they are not appropriate for cultural heritage information, at least in the way they are defined by the INSPIRE PS Data Specification. Such is the case of attributes that refer to habitats, vulnerability, the Natura 2000 network management or even human activities affecting the site's conservation.</p>
<p>TimePeriod: an attribute intended to be useful for this kind of cultural heritage data, but whose use is often problematic. An alternative chronology attribute is proposed in the Cultural Heritage Application Schema (see Chronology), for it is an attribute of the real-world entities that are included in protected sites, not of the legal object constituting a protected site. Besides, a legally defined protected site can contain several entities of different chronologies.</p>
<p>Versioning of spatial objects: it is accomplished through the use of <i>lifeCycleInfo</i> characteristic in the <i>beginLifeSpanVersion</i> and <i>endLifeSpanVersion</i> attributes.</p>

The Cultural Heritage Application Schema proposed complements the Protected Sites Full schema in two ways. First, it adds a few attributes to some of the already existing classes in the Protected Sites Full schema. Second, it incorporates a hierarchy of new classes that couple with the already existing ones, thereby allowing the incorporation of heritage data peculiarities. Most of this will be described in detail in the following sections.

3.2. The Cultural Heritage Application Schema

This proposal takes into consideration three main aspects regarding common cultural heritage management practice: the legal protection framework, the cultural entities under protection and the documentation associated to them.

3.2.1. Interoperability Features

A key goal is that all these aspects may be developed by data providers, which requires an interoperability framework.

From a legal point of view the INSPIRE PS Data Specification has been expanded to ensure a standardized inclusion of cultural heritage within the Directive, and international (as well as national for the Spanish case) specific protection figures have been added. Cultural and documentary parts of the schema are designed according to existing standards in both fields, the CIDOC-CRM covering the cultural part, and the Dublin Core Metadata Initiative (DCMI) regarding documentation.

The scope of CIDOC-CRM is much broader than the scope of the Cultural Heritage Application Schema (see section 2), for the latter refers only to certain as-

pects of cultural heritage, but some of the classes defined by the CIDOC-CRM are relevant to describe cultural heritage spatial data and, therefore, find a place in the model.

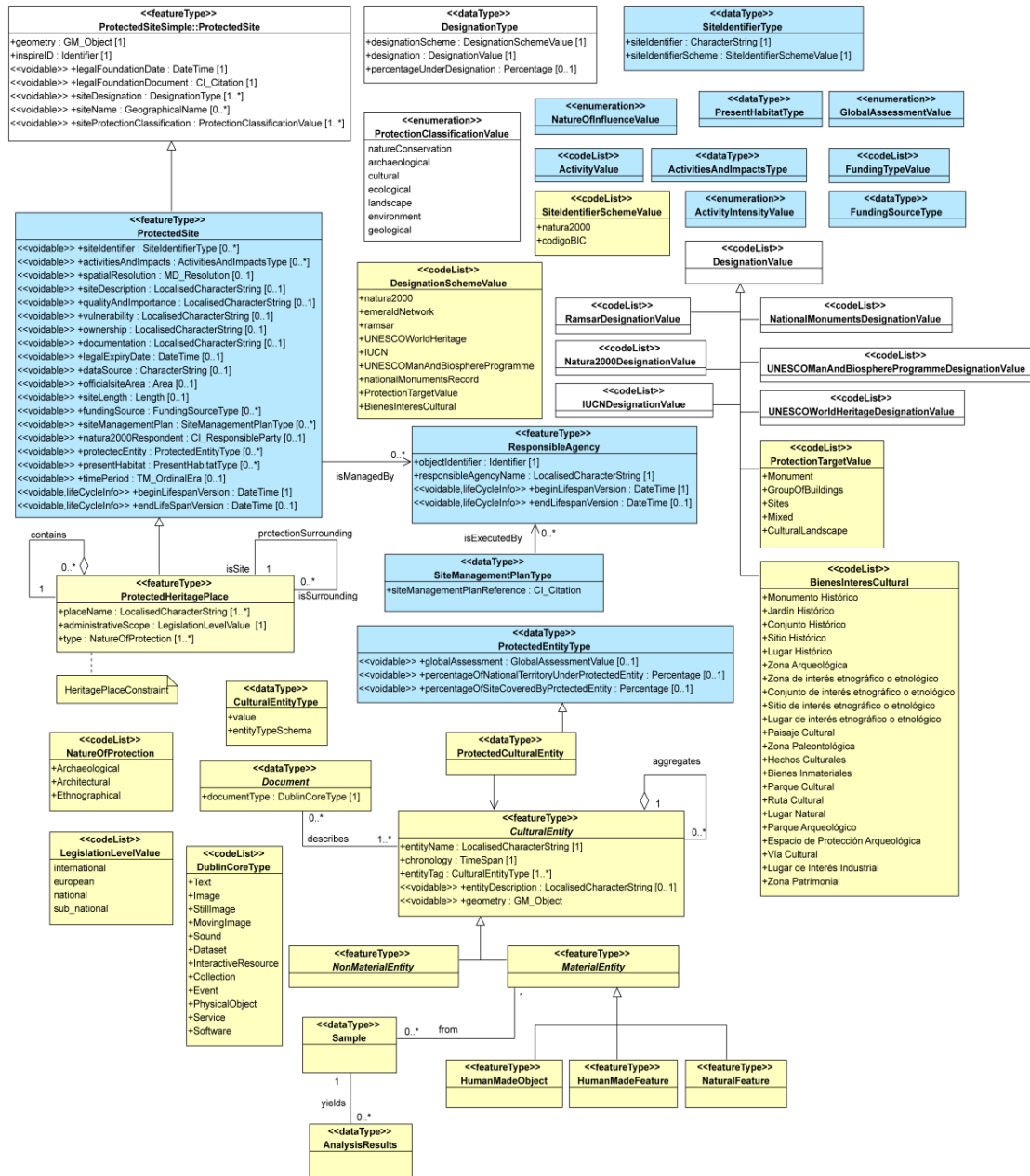
The DCMI is also a well known ISO norm (ISO 15836:2009) that establishes a standard for documentary resources description based on twelve types of documents. An attribute based on these basic categories is included to organise that information.

3.2.2. *Model Description*

The Cultural Heritage Application Schema (Figure 1) extends the Protected Sites Full Application Schema in three ways:

- It adds new code lists for the *siteDesignation* attribute. These code lists enable a generic description of the reasons that motivated the establishment of legal protection in the first place, based on both UNESCO descriptors (*ProtectionTargetValue*), and a list of Spanish protection categories (*BienesInteresCultural*).
- It adds a few attributes which are relevant to cultural heritage data through the addition of the *ProtectedHeritagePlace* class, a subtype of *ProtectedSite*. This extension also applies the constraints of *Protected sites* in order to enable the development of a common framework for cultural heritage data.
- Finally, it complements the schema by adding a group of classes to describe the real-world entities protected by the legal figure. The main class is *CulturalEntity*, that relates to the *protectedEntity* attribute of *ProtectedSites*, in the same way that it relates to other entities protected by their natural value, such as *Habitats and Biotopes*, *Species Distribution* or *Buildings*, being developed within Annex III of the INSPIRE Directive.

Figure 1: The Cultural Heritage Application Schema: General Outline



The class that gathers information on the legal protection of a site is *Protected-HeritagePlace*, which inherits the *ProtectedSite* class attributes, constraining their use through certain categories, in a small effort to offer minimal common infor-

mation (Figure 2). This constraint provides some specifications on how to fill crucial attributes.

Some of the difficulties introduced by the adaptation of cultural heritage data to the PS Data Specification are addressed with the aid of this constraint.

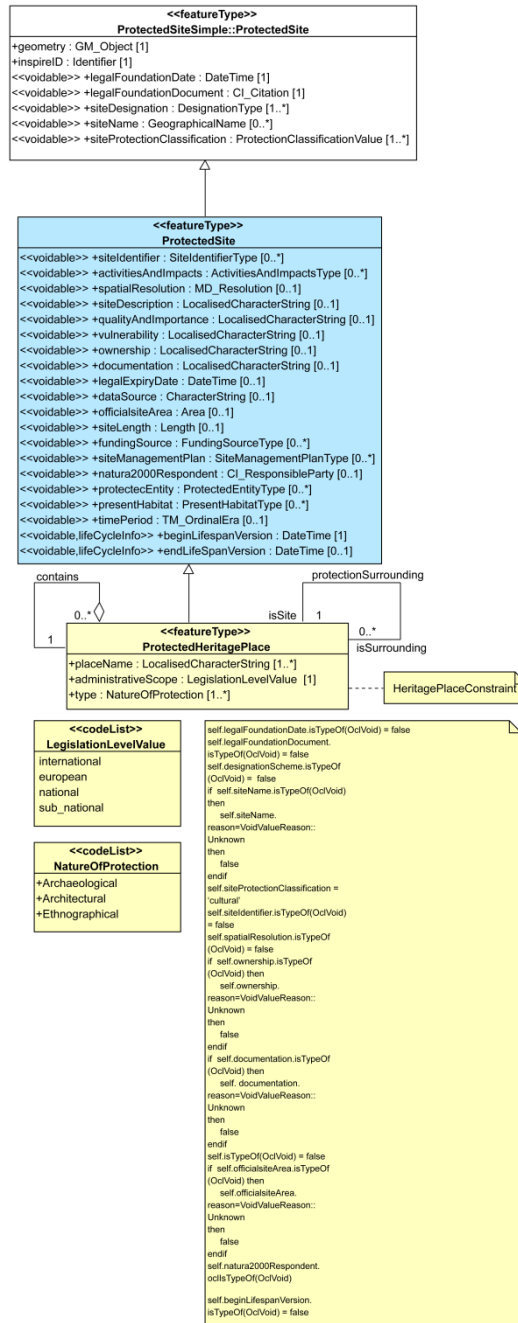
To enable a homogeneous integration of cultural heritage data, a series of steps should be undertaken for the *ProtectedSite* class:

- Voidable attributes defined in the simple schema characterize the main features of protected sites, and therefore, their voidable stereotype is removed in the Cultural Heritage Application Schema. The only exception is the site name, which has to be filled with a geographical name (from a national gazetteer) which may, however, not always exist for cultural heritage places. The attribute *siteProtectionClassification* offers the reasons behind the protection, using a code list that contains two possible values for cultural heritage, *cultural* and *archaeological*. Cultural heritage data providers will find this rather problematic, for *archaeological* is better understood as a subcategory of *cultural*, at the same level as *ethnological* or *architectural*. It is more convenient, therefore, to always use the code *cultural* at this level. Further specifications are made below.
- *siteIdentifier*, *spatialResolution*, *dataSource* and *beginLifeSpanVersion* attributes are also considered crucial in order to constitute a coherent dataset, enforcing data providers to fill them by eliminating the voidable stereotype in the constraint.
- Attributes that refer to habitats, vulnerability, the Natura 2000 network management or even to human activities that have an impact on the site's conservation are relevant only to natural protected sites, at least in the way they are defined in the PS Data Specification, Although they might be filled in for some kind of heritage places, they are not adequate for cultural heritage, so they are voidable.
- Attributes concerning ownership (its value could be *private*, *public* or *private/public*), documentation, and official site area are kept as voidable, should that information be unknown.

All *ProtectedSites* attributes are inherited by the newly created class *Protected-HeritagePlace*. This class is created to allow the inclusion of three new attributes, regarded as central in order to define this specific kind of protected sites: the name of the heritage place (that may differ from the geographical name), the legislation level of protection and the nature of the protected heritage place as stated in its definition (archaeological, architectural, ethnographical or a combination of them).

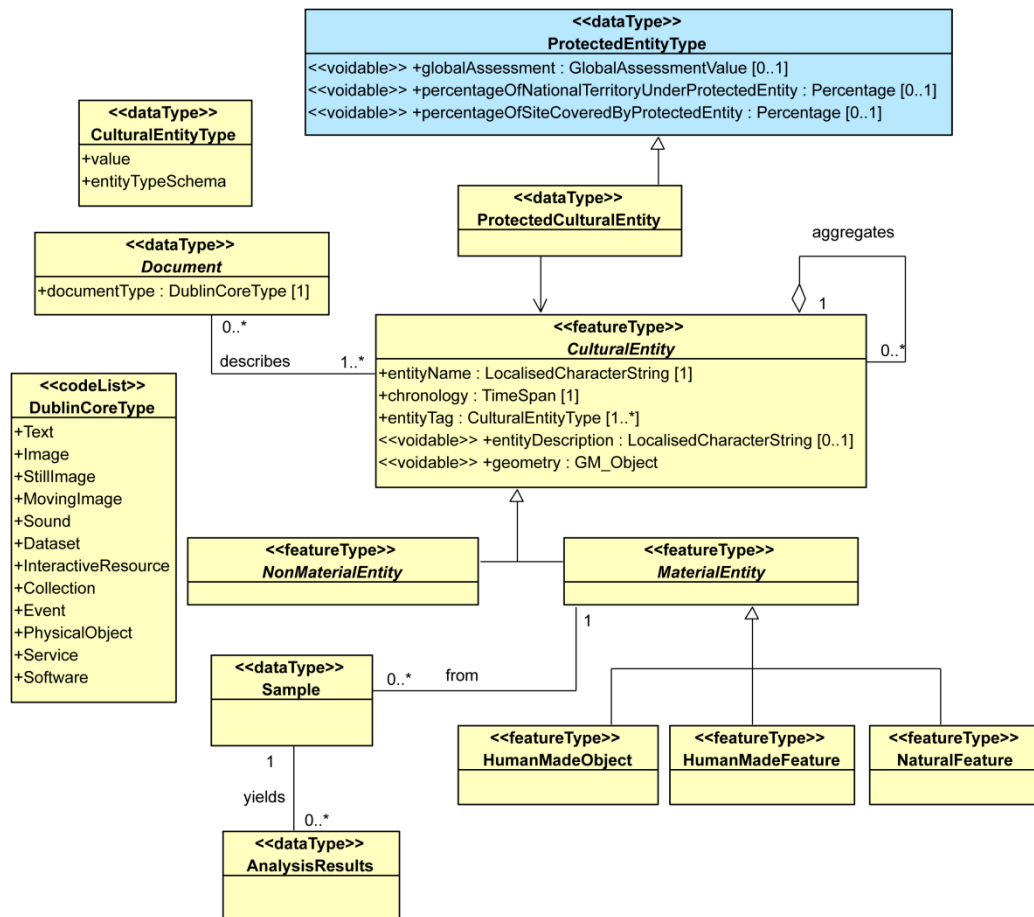
Two self-aggregation relationships have been established in order to include the usual conditions found in the documentation. The relationship named *contains* allows the inclusion of several protected sites within the legal definition of a protected heritage place, such as ethnographic, architectural or archaeological sites that may fall within the limits of a cultural landscape, for instance. The association named *protectionSurrounding* would be used when the legal document refers to the site's protected surroundings, which normally have a different legal condition and degree of protection, but whose existence is inseparably linked to the site itself. As long as the protected surroundings have different attributes, they should be instanced as another object of the same class, and referred to through the self-association named *protectionSurrounding*.

Figure 2: The Protected Heritage Place Class



The real-world entity being protected is linked to the voidable *ProtectedEntityType* class, through the intermediate class *ProtectedCulturalEntity*, designed following the model of other kinds of protected sites such as *Buildings*, *Habitats* and *Biotopes* or *Species Distribution*. The main class that systematizes all this information is *CulturalEntity* (Figure 3). This may be a large place such as a whole building or an archaeological site. But it also could be a smaller feature such as a wall or a brick, being the building an aggregation of those smaller features and a cultural entity as well. These resolution variations would be expressed by the self-aggregation relation. Consequently, the disaggregation of the elements included in the model depends on the nature of the data as gathered by the provider. Cultural entities are defined by a short set of attributes, regarding their name, description, chronology, classification and geometry.

Figure 3: Classes Related To Cultural Entities



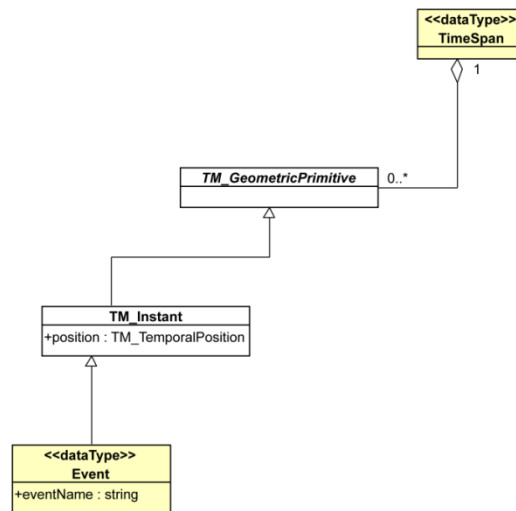
For the *chronology* attribute we rely on the implementation of ISO 19108:2002, which provides a sound schema to expose the timeline of a cultural entity, describing chronology as an element together with geometry and topology in a single dimension. A cultural entity's timeline starts at the moment of its creation and, in most cases, reaches the present day, or else a specific moment in the past. This timeline is marked by a series of events: creation, occupation, abandonment, re-occupation, modification, restoration, etc.

The schema contained in the ISO 19108:2002 suits these concepts, as it is built of geometric and topological primitives (*instants/nodes* and *periods/edges*), where instants can be defined at any degree of accuracy, depending on the Temporal Reference System used.

Reference systems have temporal and spatial extents of application and may consist of calendar dates, clock times, coordinate systems or ordinal eras. This means that they can be as accurate or flexible as needed, so data providers may implement a rich and complex chronology referencing events with a precise date, or else they may merely provide a rough indication for this attribute.

Within this rich schema, the addition of two new classes should be enough to complete the description of the timeline of cultural entities (Figure 4):

- *TimeSpan* is a collection of ISO 19108:2002 primitives that gathers the chronological timeline of a cultural entity.
- *Event* inherits from *TM_Instant*, adding a new attribute to specify the type of event recorded, which can be creation, modification, restoration, destruction, etc.

Figure 4: New temporal classes

Multiple classifications may be assigned to cultural entities under the *entityTag* attribute. The proposal of a unique thesaurus to describe such a wide variety of concepts, ranging from a “prehistoric mining site” to the “Way of St. James” is an overwhelming task that would ultimately leave too many realities aside. A more reasonable option would be to leave the selection of appropriate thesaurus – many standards already exist– to the data providers, which would have to be specified, together with the value, in the *CulturalEntityType* class.

The geometry of a cultural entity is independent of that of the protected area, which means the legal boundary of a protected site differs from the boundary of the real-world phenomenon. In practical terms, this distinction enables the inclusion of different cases that data providers will probably have to confront:

- The condition, form or geometry of a real-world object can change with no corresponding effect in the protected site that embraces it. For instance, a portion of a building may collapse, totally or partially, but the area subject to protection may remain intact for different reasons.
- The condition, form or geometry of a protected site can change without any previous modification in the real-world object related to it. For example, the protected area around a building may be extended, or reduced, for technical, legal or practical reasons, while the building that is protected remains exactly the same.

Thus, changes in attributes of the protected site (including geometry) do not imply changes in attributes of the legal entity being protected. Within the proposed

model, a cultural entity will be described and mapped just once, although it can be included in many different protected sites (such as a local urban planning, a national monuments record or a World Heritage Site).

Cultural entities are categorized according to their objective quality, so they may be material –a temple–, or immaterial –a feast associated to a fixed location–. This classification of cultural entities is based on the very basic distinction between tangible and intangible, which are concepts that refer to the physical nature of things which are considered heritage. These two categories are imminently different in how they are documented, described and spatially recorded. This is the rationale behind the subdivision of the class *CulturalEntity* into the two subclasses *NonMaterialEntity* and *MaterialEntity*. Thus, different types of things, events or manifestations that are considered relevant for the protection of a site may be separately described and spatially referenced.

Material cultural entities are arranged according to the three classes of the CIDOC-CRM that embrace all kind of material heritage phenomena: *Man Made Object*, *Man Made Feature* and *Site*, which would be translated in the Cultural Heritage Application Schema as *HumanMadeObject*, *HumanMadeFeature* and *NaturalFeature*:

- Human Made Objects are those physical features that are purposely created by human activity, and have a physical boundary that separates them from other objects (e.g. wall, building).
- Human Made Features are physically integrated inside other objects, with no clear boundaries (e.g. rock art, pit, hypogeous).
- Natural Features are those which are singularly identifiable, and endowed with a cultural significance for a community (e.g. forest, beach, mountain).

Samples may be taken from material entities, in order to determine their nature, condition, state of preservation, date, etc. The class *Sample* is designed to collect that information, along with the *AnalysisResults* class, whose purpose is to describe the different analytical results that they may yield.

Documents that are associated with a cultural entity are arranged according to the *type* attribute of the Dublin Core Element Set, which establishes twelve specific kinds of documents according to the nature of the information they contain:

- *Text*: Its distinctive feature is that it consists primarily of words for reading, though it would also include the images associated to the text as recommended by Dublin Core. Depending on the cultural entities, it can refer to scientific publications, grey literature, etc.
- *Image*: Any kind of visual representation.

- *StillImage*: It is a subtype of image, characterized by its static nature, such as a photograph or a drawing.
- *MovingImage*: It is another subtype of image that includes moving pictures, animations, videos, etc.
- *Sound*: An acoustic representation, intended to be heard, such as story recordings, music, etc.
- *Dataset*: A set of data stored or organized in a structured way, such as a list, a table or a database.
- *InteractiveResource*: A resource requiring interaction from the user to be understood, executed, or experienced, such as web pages, virtual reality objects, etc.
- *Collection*: An aggregation of resources, containing documents of any kind.
- *Software*: A computer program.
- *Service*: A system that furnishes a certain function, such as a web server.
- *PhysicalObject*: A real-world object.
- *Event*: An occurrence that happens within a certain lapse of time.

4. DISCUSSION

The proposed Cultural Heritage Application Schema provides a conceptual framework for managing and distributing cultural heritage information in SDIs in accordance with the INSPIRE Directive. Several characteristics prove the appropriateness of this objective:

- It offers conceptual interoperability, a common data model for adapting and sharing georeferenced information generated by different data providers.
- Although linked to the wider fields of geographical information and protected sites, it is specifically aimed at cultural heritage data, incorporating entities and attributes that acknowledge their idiosyncrasy.
- It distinguishes the two conceptual levels: cultural features –the real-world entities– and protected sites –administrative areas that are “artificially” created for managing and preserving cultural features–.
- It is a modular schema that allows several levels of description, from a generic indispensable package limited to one entity (*ProtectedHeritagePlace*) and thirteen attributes, to a more complex set that allows adding other kinds of information if desired. It is also extendable through the addition of new subclasses, with the aim of incorporating more specific types of cultural heritage information, concerning particular disciplines, thematic areas, datasets, etc.

In this way, the whole range of agents involved in cultural heritage may benefit from a common framework of data distribution:

- Heritage agencies, which can benefit from a transparent and precise platform for the control of heritage assets, and a means to enhance its dissemination to society in a digital world.
- Administrations, benefiting from improved coordination of their management policies by allowing a better integration between cultural heritage sites and other relevant geographical features (e.g. geology, public works, equipments and services, etc.).
- Research institutions, both as cultural heritage data consumers and producers, encouraging the publication of their datasets via distributed information systems (such as SDIs).
- Finally, the general public, by having a better access to heritage information, in accordance with the legal directives which recognise the citizen rights of modern democratic states.

The Cultural Heritage Application Schema is currently being experimented in some archaeological SDIs: IDEArq, a SDI planned within the Spanish National Research Council (CSIC) and focused on the dissemination of archaeological project data; and IDEPatri (IDE Arqueológica da Idade de Ferro en Galicia), an SDI for the publication of descriptive and analytical information available regarding the Iron Age sites of the NW of the Iberian Peninsula.

We hope that this proposal will be helpful for the future development of cultural heritage SDIs in an interoperable framework based on OGC standards. These, by taking into account the spatial nature of this specific kind of protected sites data, will enhance their role within territorial governance, assisting in their management and protection, and to bring them closer to the general public.

ACKNOWLEDGEMENTS

As was abovementioned, the initial steps of the work presented here were taken within a Thematic Sub-group of the Spanish SDI Working Group. The authors wish to acknowledge the following members of that sub-group for their contribution (in alphabetical order): Tomás Abad Balboa, Emilio Abad Vidal, Pilar Chías Navarro, Marta Criado Valdés, Mercedes Farjas Abadía, Francisco García Cepeda, Miguel Lage Reis-Correia, Javier Márquez Piqueras, Victorino Mayoral Herrera, Juan Luis Pecharromán Fuente, Arantza Respaldiza Hidalgo, María Ruiz del Árbol Moro, Antonio Vázquez Hoenhe and José Julio Zancajo Jiménez.

We would also like to thank Sebastián Mas Mayoral and Antonio Rodríguez Pascual for the possibility of developing the work of the GTT-PAH inside the Spanish SDI Working Group (GT-IDEE).

The work presented here has been partially supported by the Spanish Government through the projects “Investigación en tecnologías para la valoración y con-

servación del Patrimonio Cultural” (CONSOLIDER INGENIO 2010 CSD2007-00058) and SPATRIAL-“Definición de un Modelo para la Representación de Aspectos Espaciales del Patrimonio” (HAR2009-1226, Programa Nacional de I+D+I, Ministry of Science).

REFERENCES

- Corns, A.J. and R. Shaw (2010). “Cultural Heritage Spatial Data Infrastructures (SDI) - Unlocking the Potential of our Cultural Landscape Data”, in Reuter, R. (Ed.) *30th EARSeL Symposium Remote Sensing for Science, Education, and Natural and Cultural Heritage*, pp. 1–8.
- Cripps, P., Greenhalgh, A., Fellows, D., May, K. and D. Robinson (2004). *Ontological Modelling of the Work of the Centre for Archaeology*, CIDOC CRM Technical Paper.
- Crofts, N., Doerr M., Gill T., Stead, S. and M. Stiff. (2010). *Definition of the CIDOC Conceptual Reference Model*, ICOM/CIDOC CRM Special Interest Group.
- Doerr, M. (2003). The CIDOC Conceptual Reference Module. An Ontological Approach to Semantic Interoperability of Metadata. *AI Magazine* 24 (3): 75–92.
- English Heritage (2012). MIDAS Heritage - the UK Historic Environment Data Standard. http://www.english-heritage.org.uk/publications/midas-heritage/midas-heritage-2012-v1_1.pdf [accessed 14 January 2014].
- Faro (2005). The Faro Convention on the Value of Cultural Heritage for Society. Council of Europe. at <http://conventions.coe.int/Treaty/EN/Treaties/Html/199.htm> [accessed 14 January 2014].
- Hiebel, G., Hanke, K. and I. Hayek (2013). “Methodology for CIDOC CRM Based Data Integration with Spatial Data”, in Contreras, F., Farjas, M. and F. J. Melero (Eds). *CAA 2010 Proceedings*. Granada, pp. 547–554.
- IAPH (2011). *Recomendaciones Técnicas Sobre La Georreferenciación de Entidades Patrimoniales*. at http://www.iaph.es/export/sites/default/galerias/patrimonio-cultural/documentos/gestion-informacion/RTCI_6_georreferenciacion_v1.3.pdf / [Accessed 14 January 2014].

- IDE Arqueolóxica da Idade de Ferro en Galicia at <http://idepatri.cesga.es/> [Accessed 20 June 2012].
- Directive 2007/2/EC of the Council of the European Union and European Parliament Setting up an Infrastructure for Spatial Information in the European Community (INSPIRE).
- INSPIRE Thematic Working Group Protected sites (2010). *INSPIRE Data Specification on Protected Sites – Guidelines v3.1*.
- Janowicz K. (2009). "The Role of Place for the Spatial Referencing of Heritage Data," presented at *The Cultural Heritage of Historic European Cities and Public Participatory GIS Workshop*, The University of York, UK.
- Katsianis, M., Tshipidis, S., Kotsakis, K. and A. Kousoulakou (2008). A 3D Digital Workflow for Archaeological Intra-Site Research Using GIS, *Journal of Archaeological Science* 35 (3): 655–667.
- Lage, M., Corns, A., Pecharromán, J.L. and R. Shaw (2009). "Integration of European Cultural Landscapes Data within Spatial Data Infrastructures (SDI)", in del Valle, A. O. S., Manttingly, D. and M. Clavel-Lévêque (Eds.), *From Present to Past Through Landscape*. Madrid: CSIC, pp. 221–228.
- McKeague, P. (2012). "Historic Environment and INSPIRE -- A View from Scotland", in M. Ioannides, D. Fritsch, and J. Leissneret (Ed). *Progress in Cultural Heritage Preservation. 4th International Conference, EuroMed 2012 Limassol, Cyprus, October 29 – November 3, 2012. Proceedings*. Springer – Springer Lecture Notes in Computer Science-, pp. 833–840.
- McKeague, P., Corns, A. J. and R. Shaw (2012). Developing a Spatial Data Infrastructure for Archaeological and Built Heritage, *International Journal of Spatial Data Infrastructures Research*, 7: 38–65.
- Meyer, E., Grussenmeyer, P., Perrin, J. P., Durrand, A. and P. Drap (2007). A Web Information System for the Management and the Dissemination of Cultural Heritage Data, *Journal of Cultural Heritage* 8 (4): 396–411.
- Omtzigt, N., Reus, N. D., Koomen, E. and A. Arezza (2008). "The Construction of a Spatial Data Infrastructure for Cultural Heritage", in Cunningham, P. and M. Cunningham (Eds). *Collaboration and the Knowledge Economy: Issues, Applications, Case Studies*. Amsterdam: IOS Press, pp. 1366–1373.

- Papatheodorou, C., Carlisle, P., Ertmann-Christiansen, C. and K. Fernie (2011). *The CARARE Metadata Schema, v.1.1*.
- Parcero-Oubiña, C. (2012). "Heritage-Oriented Spatial Data Infrastructures in Spain: Waiting on the World to Change", in Ioannides, M., Fritsch, D. and J. Leissneret (Eds). *Progress in Cultural Heritage Preservation. 4th International Conference, EuroMed 2012 Limassol, Cyprus, October 29 – November 3, 2012. Proceedings*. Springer – Springer Lecture Notes in Computer Science-, pp.692–700.
- Pfoser, D., Hadzilacos, T., Faradouris, N. and K. Kyrimis (2007). "Spatial Data Management Aspects in Archaeological Excavation Documentation." In Fabrikant, S. and M. Wachowicz (Eds), *The European Information Society*, Berlin, Heidelberg: Springer, pp. 287–302.
- Querol, M.A. (2010). *Manual De Gestión Del Patrimonio Cultural*. Madrid: Akal.
- Strle, G. and M. Marolt (2012). The EthnoMuse Digital Library: Conceptual Representation and Annotation of Ethnomusicological Materials. *International Journal on Digital Libraries* 12 (2-3): 105–119.
- Thornes, R., and J. Bold (Eds). (1998). *Documenting the Cultural Heritage*. Getty Information Institute.
- Tudhope, D., May, K., Binding, C. and A. Vlachidis (2011). Connecting Archaeological Data and Grey Literature via Semantic Cross Search. *Internet Archaeology* 30 at <http://intarch.ac.uk/journal/issue30/5/1.html>, [accessed 14 January 2014].
- Vicent García, J.M. (2007). "La Arqueología a comienzos del siglo XXI: Ciencia, tecnología, valores y sociedad". *Actas de XII Jornadas de Estudios sobre Lanzarote y Fuerteventura*. Arrecife, Lanzarote: Cabildo Insular de Lanzarote, pp. 327–347.
- Willmes, C, Brocks, S., Hoffmeister, D., Hütt, C., Kürner, D., Volland, K. and G. Bareth (2012). Facilitating Integrated Spatio-Temporal Visualization and Analysis of Heterogeneous Archaeological and Palaeoenvironmental Research Data, *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* I (2): 223–228.